

Regina Bernal^I

Nilza Nunes da Silva^{II}

Cobertura de linhas telefônicas residenciais e vícios potenciais em estudos epidemiológicos

Home landline telephone coverage and potential bias in epidemiological surveys

RESUMO

OBJETIVO: Estimar o efeito da taxa de cobertura de linhas telefônicas residenciais em potenciais vícios de informação em inquéritos epidemiológicos.

MÉTODOS: Foram utilizadas as bases de dados da Pesquisa Nacional por Amostra de Domicílios no período de 1998 a 2003 para a estimativa das taxas de cobertura de linhas telefônicas residenciais nas cinco regiões geográficas brasileiras. Utilizou-se a regressão logística múltipla para identificar os fatores associados à posse de linha telefônica fixa. O impacto do vício nos intervalos com 95% de confiança foi avaliado em função da precisão alcançada em cada situação.

RESULTADOS: Nas regiões metropolitanas Sudeste, Sul e Centro-Oeste com 70% e mais de cobertura, os vícios associados foram considerados desprezíveis. Nas demais regiões, os vícios relativos estavam acima do limite aceitável (0,4), indicando possíveis erros nas inferências construídas sob intervalo com 95% de confiança. A chance de acesso à linha telefônica residencial foi maior para população com cor da pele branca e maior escolaridade.

CONCLUSÕES: Os achados mostram que o uso de cadastro de linhas telefônicas residenciais é indicado para a realização de inquéritos epidemiológicos apenas para estados com cobertura acima de 70%. Metodologias específicas para o tratamento de estimativas obtidas em localidades com taxas inferiores, precisam ser estudadas e divulgadas.

DESCRIPTORIOS: Levantamentos Epidemiológicos. Vícios de cobertura. Vigitel. Entrevistas por telefone.

^I Programa de Pós-Graduação em Saúde Pública. Faculdade de Saúde Pública. Universidade de São Paulo (FSP-USP). São Paulo, SP, Brasil

^{II} Departamento de Epidemiologia. FSP-USP. São Paulo, SP, Brasil

Correspondência | Correspondence:

Regina Tomie Ivata Bernal
R. Geraldo Amorim, 257
05594-110 São Paulo, SP, Brasil
E-mail: rbernal@usp.br

Recebido: 20/04/2008
Revisado: 22/08/2008
Aprovado: 23/09/2008

ABSTRACT

OBJECTIVE: To estimate landline telephone coverage effects on potential information bias in epidemiological surveys.

METHODS: Databases of the *Pesquisa Nacional por Amostra de Domicílios* (National Household Sample Survey), from 1998 to 2003, were used to estimate landline telephone coverage rates in the five Brazilian geographic regions. Multiple regression analysis was used to identify factors associated with owning a landline telephone. Bias impact on 95% confidence intervals was assessed, according to the accuracy achieved in each situation.

RESULTS: In the Southern and Central-west metropolitan regions, with 70% coverage or more, associated bias was considered insignificant. In the remaining regions, related bias was above the acceptable limit (0.4), indicating possible errors in inferences drawn with a 95% confidence interval. Likelihood of access to a home landline telephone was greater among the white population with higher level of education.

CONCLUSIONS: Findings reveal that the use of residential telephone records is indicated to conduct epidemiological surveys in Brazilian states where coverage is above 70% exclusively. Specific methodologies to treat estimates obtained in regions with lower rates need to be analyzed and published.

DESCRIPTORS: Epidemiological Surveys. Coverage bias. Vigitel. Telephone Surveys.

INTRODUÇÃO

No Brasil, é crescente o interesse de pesquisadores brasileiros no uso de linhas telefônicas residenciais (LTR) para realização de inquéritos epidemiológicos. A produção científica existente expõe a definição do cadastro como uma das principais questões relacionadas ao plano de amostragem.^{2,3,4,a} Em geral, os inquéritos com base populacional utilizam as listas telefônicas de assinante para sortear amostras de domicílios que possuem telefones residenciais. Entre os problemas decorrentes desta prática está a desatualização das informações disponíveis, exigindo, na maioria dos trabalhos, a necessidade de compensá-la por meio da busca de telefones via internet, disponíveis no site da empresa telefônica responsável. Contudo, as taxas de resposta alcançadas (acima de 60%) estimulam o uso dessa estratégia para a realização de inquéritos mais ágeis e econômicos, proporcionando redução de custos e prazos de realização de entrevistas.

No Brasil, a Secretaria de Vigilância em Saúde do Ministério da Saúde (SVS/MS) vem desenvolvendo o Sistema de Vigilância de Fatores de Risco para Doenças Crônicas por Inquérito Telefônico (VIGITEL)

desde 2003.^{5,b} Esta pesquisa é realizada anualmente, o que permite acompanhar a evolução dos hábitos e comportamentos da população brasileira. No entanto, há necessidade de examinar metodologias específicas para as questões como: cobertura, taxa de resposta e eficiência (relativa aos pesos de pós-estratificação para corrigir vícios nas estimativas das variáveis de interesse).

O objetivo do presente estudo foi estimar o efeito da taxa de cobertura de linhas telefônicas residenciais em potenciais vícios de informação em inquéritos epidemiológicos.

MÉTODOS

Para estimar as taxas de cobertura de domicílios com LTR no Brasil foram utilizadas as bases de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE), no período de 1998-9 e 2001-3.^c Essa pesquisa é realizada anualmente pelo processo de sorteio de amostragem probabilística de domicílios com abrangência nacional, exceto a área rural da região Norte. A população-alvo

^a Scheuren F. What is a survey? American Statistical Association. Washington D.C.; 2004 [citado 2009 mar 20]. Disponível em: <http://www.whatisasurvey.info>

^b Ministério da Saúde. Vigitel Brasil 2006: Vigilância de fatores de risco e proteção de doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sócio-demográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2006. Brasília; 2007.

^c Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

é composta pelos domicílios e pessoas residentes em domicílios nas áreas pesquisadas. A PNAD adota o Plano Complexo de Amostragem (PCA) que envolve a estratificação geográfica de conglomerados com um, dois ou três estágios de seleção.⁷ O território brasileiro foi dividido em 36 estratos, constituídos por: 17 nas unidades da federação (RO, AC, AM, RR, AP, TO, MA, PI, RN, PB, AL, SE, ES, SC, MT, MS, GO) e o Distrito Federal, definidos como estratos independentes; e 18 nas demais unidades da federação (PA, CE, PE, BA, MG, RJ, SP, PR e RS) formaram-se estratos: um com região metropolitana e outro com os municípios restantes. Foram consideradas nas análises as variáveis que definiam a estrutura do plano amostral, denominadas estrato, a unidade primária de amostragem (UPA) e o peso amostral. As análises estatísticas foram realizadas com o *software* Stata versão 8.0, do qual se utilizou o módulo “svy” adequado para obtenção de estimativas não viciadas quando os dados são provenientes de planos de amostragem complexos.

A taxa de cobertura de LTR para cada ano fixado foi

$$\text{estimada por uso da expressão } r = \frac{\sum_h \sum_\alpha \sum_\beta w_h y_{h\alpha\beta}}{\sum_h \sum_\alpha \sum_\beta w_h x_{h\alpha\beta}}$$

onde: $y_{h\alpha\beta}$ é o valor observado da variável (1 se possui telefone fixo, e 0 se não) do domicílio β , da UPA α ; do estrato h ; $x_{h\alpha\beta}$ igual a 1, para todos os domicílios β observados na UPA α , do estrato h ; $w_h = \frac{1}{f_h}$, peso de cada domicílio é dado pelo inverso da fração de amostragem de cada estrato h .

Os intervalos de confiança para taxas de cobertura foram estimados pela expressão $r \pm t_{0,95,d} [\text{var}(r)]^{1/2}$, sendo $\text{var}(r) = \frac{1}{x^2} [\text{var}(y) - 2r \text{cov}(y, x) + r^2 \text{var}(x)]$.⁴

Quanto ao perfil de usuário de LTR, em cada região geográfica, os entrevistados com 18 anos ou mais foram previamente divididos em dois grupos mutuamente exclusivos, contendo aqueles que residiam em domicílios com ou sem LTR. Mediante uso do modelo de regressão logística múltipla^a

$$\log \left\{ \frac{\pi(x)}{1 - \pi(x)} \right\} = \beta_1 + \beta_2 x_2 + \dots + \beta_p x_p, \text{ em que } \pi(x)$$

é a probabilidade de acesso a LTR em função das variáveis (x_i), foram definidas oito características sociodemográficas (Tabela 1). O critério da inclusão final dessas variáveis no modelo seguiu a técnica *stepwise*, considerando o nível de significância de 5%.

Para estimar o vício potencial associado à não cobertura de LTR, os dados da PNAD 2003, a qual trouxe como tema complementar as características de saúde, foram

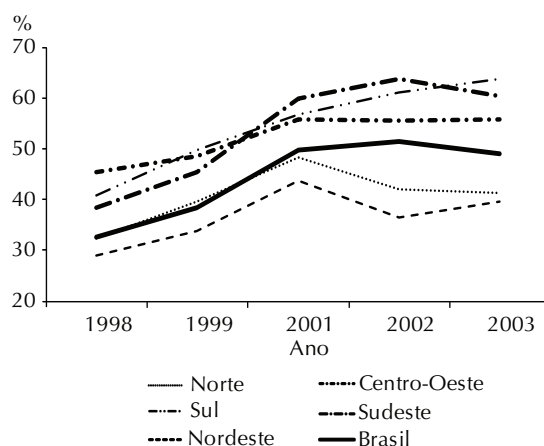
usados para obter estimativas da prevalência de doenças crônicas para a população maior de 40 anos residentes nas áreas estudadas. Assim, os entrevistados foram separados em dois estratos: os que possuem e os que não possuem LTR. Foram considerados portadores de doenças crônicas os indivíduos que confirmaram o diagnóstico, referido por médico ou profissional de saúde, de pelo menos uma das doenças entre: doença da coluna ou das costas, artrite ou reumatismo, diabetes, bronquite ou asma, doença do coração ou hipertensão.

O vício potencial devido a baixas taxas de cobertura foi estimado pela expressão $Vício\ Relativo = \frac{|V|}{EP(P)}$,^{3,6}

em que o numerador é o valor absoluto calculado para o vício $V = \frac{N_2}{N} (P_1 - P_2)$ decorrente da diferença entre as prevalências dos dois estratos, ponderada pela proporção de pessoas residentes em domicílios sem acesso a LTR que formaram o segundo estrato. Valores acima de 0,40 para o vício relativo (VR) alteram o nível de 95% inicialmente fixado para os intervalos de confiança (Cochran 1997),¹ podendo levar à construção de inferências equivocadas.

RESULTADOS

Os resultados alcançados mostraram crescimento de 50% dos domicílios cobertos por LTR no período de 1998 a 2003 no Brasil (Figura 1). No entanto, essa evolução não ocorreu de maneira uniforme no País. Nas regiões Norte, Nordeste e Centro-Oeste, as proporções de linhas telefônicas residenciais cresceram 26%, 36% e 22%, respectivamente, enquanto nas



Fonte: Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

Figura 1. Evolução da cobertura de domicílios com linhas telefônicas residenciais por região. Brasil, 1998-2003.

^a Paula GA. Modelos de Regressão com apoio computacional [tese de doutorado]. São Paulo: Instituto de Matemática e Estatística da USP; 2004.

Tabela 1. Características sociodemográficas dos adultos residentes em domicílios segundo posse de linha telefônica residencial. Brasil, 1998-2003.

Variável	Posse de linha telefônica residencial			
	Não		Sim	
	%	Erro-padrão	%	Erro-padrão
Anos de estudo				
0 - 4	69,0	0,4	31,0	0,4
5 - 8	55,3	0,4	44,7	0,4
9 - 11	44,4	0,4	55,6	0,4
12 - 15	26,5	0,3	73,5	0,3
15 -	7,4	0,3	92,6	0,3
Sexo				
Masculino	48,5	0,3	51,5	0,3
Feminino	44,5	0,3	55,5	0,3
Cor				
Não Branca	54,9	0,4	45,1	0,4
Branca	36,1	0,4	63,9	0,4
Faixa etária (anos)				
18 - 40	50,6	0,3	49,4	0,3
41 - 60	39,9	0,4	60,1	0,4
60 -	42,7	0,5	57,3	0,5
Região metropolitana				
Não	56,2	0,5	43,8	0,5
Sim	37,0	0,4	63,0	0,4
Região				
Norte Urbana	54,4	0,9	45,6	0,9
Nordeste	56,4	0,6	43,6	0,6
Sudeste	36,2	0,5	63,8	0,5
Sul	33,4	0,8	66,6	0,8
Centro-Oeste	39,9	1,0	60,1	1,0
Área				
Rural	90,9	0,7	9,1	0,7
Urbana	41,1	0,3	58,9	0,3
Posse plano de saúde				
Não	56,9	0,3	43,1	0,3
Sim	16,8	0,3	83,2	0,3

Fonte: Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

regiões Sudeste e Sul cresceram 58% e 56%, respectivamente. Apesar desse aumento, em 2003, a taxa média de domicílios brasileiros cobertos por LTR foi de 50,3%, registrando-se que as regiões Norte (41%) e Nordeste (39%) estiveram abaixo dessa média, ao contrário das regiões Sudeste (60%), Sul (63%) e Centro-Oeste (55%).

Além disso, a proporção de adultos residentes em domicílios sem LTR foi maior nas áreas rurais e nas regiões não metropolitanas (Tabela 1). Os adultos declarados

como tendo cor da pele não-branca apresentaram maior proporção de não-acesso à LTR do que os de cor da pele branca. A quantidade de adultos sem LTR foi inversamente proporcional aos anos de escolaridade.

As estimativas obtidas pelo modelo de regressão ajustado apontaram associações entre probabilidade de acesso a LTR, características sociodemográficas e região geográfica (Tabela 2). A constante do modelo negativa significou probabilidade muito baixa de o indivíduo possuir LTR. Entretanto, fatores como escolaridade, cor, posse de pelo menos um plano de saúde, região metropolitana e área urbana ou rural do domicílio estiveram positivamente associados com o acesso a LTR. A Figura 2 apresenta os gráficos das probabilidades de possuir uma LTR segundo as variáveis cor da pele e escolaridade. Observou-se que a probabilidade de possuir LTR entre adultos com cor da pele branca e menos de cinco anos de estudo foi inferior a 50%, enquanto os adultos declarados como não-brancos necessitaria de pelo menos dez anos de estudo para alcançarem esta mesma probabilidade de acesso a LTR.

As estimativas de prevalências de doenças crônicas, para indivíduos com 40 anos e mais, apresentaram vícios relativos considerados não-desprezíveis. Os vícios relativos introduzidos por baixas taxas de cobertura segundo região mostraram diferentes comportamentos, apresentados na Tabela 3. As regiões Norte, não metropolitanas do Centro-Oeste, do Nordeste e do Sudeste, com taxas de cobertura de LTR abaixo de 60%, apresentaram vícios relativos acima de 1, o que reduz o nível de confiança pré-fixado (95%) para 83%. Na região metropolitana do Nordeste, registrou-se alteração para aproximadamente 90%. Nas regiões metropolitanas do Sudeste, Sul e do Centro-Oeste, com taxas de cobertura acima de 70%, os vícios associados foram considerados desprezíveis.

DISCUSSÃO

No período estudado, registrou-se evolução crescente das taxas de cobertura de linhas telefônicas residenciais no Brasil, com maiores índices para as regiões metropolitanas das regiões Sul e Sudeste, mas atingindo cobertura média inferior a 60% em 2003. Dessa maneira, o Brasil encontra-se inferiormente distante das coberturas praticadas nos países desenvolvidos que usam linhas telefônicas residenciais para realizar inquéritos epidemiológicos.^{2,3}

Problemas relacionados à exclusão de domicílios sem acesso a linha telefônica são objeto de estudo de publicações e de reuniões científicas internacionais voltadas à área de metodologia estatística desde 1987. Segundo Groves et al² (2001), erros devidos à exclusão de domicílios sem linha telefônica ou com número do telefone não disponível no cadastro da empresa não são desprezíveis. Para compensar a exclusão dessa

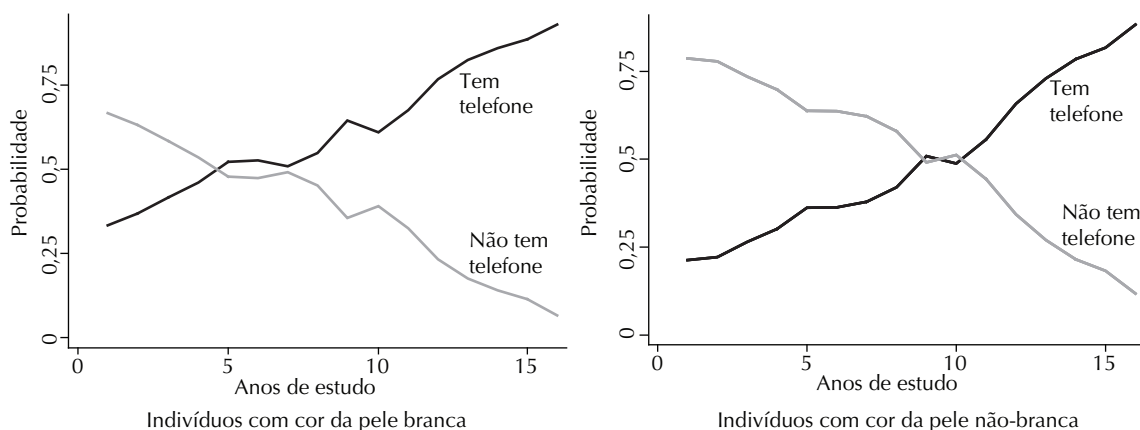
Tabela 2. Estimativas dos parâmetros associados ao modelo logístico por região do país. Brasil, 1998-2003.

Região	Variável	Estimativa	Erro-padrão	t	P> t
Norte	Constante	-1,355	0,056	-24,250	0,000
	Anos estudo	0,105	0,005	20,190	0,000
	Cor branca	0,316	0,041	7,640	0,000
	Tem plano saúde	1,270	0,057	22,090	0,000
Nordeste	Constante	-4,106	0,147	-27,950	0,000
	Anos estudo	0,139	0,003	44,290	0,000
	Cor branca	0,189	0,030	6,310	0,000
	Tem plano saúde	1,430	0,041	34,800	0,000
	Área urbana	2,691	0,149	18,070	0,000
Sudeste	Constante	-2,956	0,125	-23,580	0,000
	Anos estudo	0,112	0,031	36,250	0,000
	Cor branca	0,391	0,029	13,620	0,000
	Tem plano saúde	1,293	0,035	36,830	0,000
	Região metropolitana	0,445	0,043	10,380	0,000
	Área urbana	2,061	0,127	16,280	0,000
Sul	Constante	-2,581	0,137	-18,520	0,000
	Anos estudo	0,133	0,005	25,110	0,000
	Cor branca	0,383	0,051	7,510	0,000
	Tem plano saúde	1,061	0,046	22,980	0,000
	Região metropolitana	0,179	0,061	2,950	0,000
	Área urbana	1,824	0,127	14,410	0,000
Centro-Oeste	Constante	-3,189	0,199	-15,960	0,000
	Anos estudo	0,109	0,005	21,560	0,000
	Cor branca	0,360	0,041	8,870	0,000
	Tem plano saúde	1,153	0,061	18,810	0,000
	Região metropolitana	0,730	0,087	8,350	0,000
	Área urbana	2,319	0,205	11,290	0,000

Fonte: Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

população, tornou-se usual ajustar a distribuição da amostra com telefone à composição da população total segundo características associadas à posse de linha

telefônica por meio de processos de ponderação. Dessa maneira, espera-se reduzir vícios potencialmente associados às inferências construídas. No presente estudo,



Fonte: Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

Figura 2. Probabilidade de posse de linha telefônica residencial segundo cor da pele e anos de estudo. Brasil, 1998-2003.

Tabela 3. Taxas de cobertura e vícios relativos de prevalências estimadas para doenças crônicas entre adultos com 40 anos e mais, segundo região. Brasil, 1998-2003.

Região	Taxa de cobertura de linhas telefônicas (%)	Vício relativo
RM Centro-Oeste	85,2	0,2110
RM Sudeste	77,6	0,0751
RM Sul	76,3	0,2363
RM Nordeste	62,9	0,5598
NRM Sul	59,7	0,0172
NRM Sudeste	59,7	1,6372
NRM Centro-Oeste	56,0	1,0456
Norte	47,7	1,0552
NRM Nordeste	28,0	2,3365

RM: Região metropolitana

NRM: Não região metropolitana

Fonte: Instituto Brasileiro de Geografia e Estatística. Microdados da Pesquisa Nacional por Amostra de Domicílios-1998 a 2003. Rio de Janeiro; 2003.

após os procedimentos de ponderação das proporções, observou-se que, na região Norte urbana, escolaridade, cor da pele e posse de plano de saúde do chefe do domicílio foram associadas à posse de linha telefônica. Na região Nordeste foram observadas as mesmas características anteriormente citadas e a variável área urbana. Nas regiões Sudeste, Sul e Centro-Oeste, todas essas características estiveram presentes, mas diferenciaram-se pela variável região metropolitana.

REFERÊNCIAS

1. Cochran WG. Sampling Techniques. 3. ed. New York: John Wiley & Sons; 1977. p.14.
2. Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls II WL, Waksberg J. Telephone Survey Methodology. New York: John Wiley & Sons; 2001.
3. Kalton G. Introduction to Survey Sampling. Newbury Park: SAGE Publications Inc; 1987.
4. Kish L. Survey sampling. New York: John Wiley & Sons; 1965.
5. Monteiro CA, Moura EC, Jaime PC, Lucca A, Florindo AA, Figueiredo ICR, et al. Monitoramento de fatores de risco para doenças crônicas por entrevistas telefônicas. *Rev Saude Publica*. 2005;39(1):47-57. DOI: 10.1590/S0034-89102005000100007
6. Silva NN. Amostragem Probabilística. 2. ed. São Paulo: EDUSP; 2000.
7. Silva PLN, Pessoa DGC, Lila MF. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Cienc Saude Coletiva*. 2002;7(4):639-70. DOI: 10.1590/S1413-81232002000400005

Artigo baseado na dissertação de mestrado de Bernal R, apresentada à Faculdade de Saúde Pública da Universidade de São Paulo, em 2006.

No entanto, a região não metropolitana do Sul com 60% dos domicílios atendidos por linha telefônica apresentou o menor vício relativo. Esse resultado decorre das pequenas diferenças entre as prevalências de doenças crônicas estimadas para os estratos com e sem linha telefônica residencial. Esse resultado reforça a importância de que metodologias específicas para o tratamento de estimativas obtidas em localidades com índices inferiores precisam ser estudadas e divulgadas preferencialmente em periódicos científicos brasileiros.

A existência de um subgrupo da população residente em domicílio não atendido por uma linha telefônica, assim como suas características sociodemográficas e localidade geográfica é uma fonte de erro associado à exclusão desses domicílios, o que pode alterar o nível de confiança pré-fixado para as estimativas a serem alcançadas. Assim, baixas taxas de cobertura de linhas telefônicas residenciais podem levar a obtenção de estimativas viciadas não correspondentes às realidades estudadas.

Os resultados do presente estudo indicam que o uso de cadastro de linhas telefônicas residenciais para a realização das entrevistas em inquéritos por amostragem é confiável apenas para as unidades da federação com taxas de cobertura de domicílios com LTR acima de 70%.

Estudos brasileiros de metodologias alternativas para o uso de inquérito por telefone em regiões de baixa cobertura de linha telefônica residencial devem constituir objeto de pesquisas epidemiológicas.